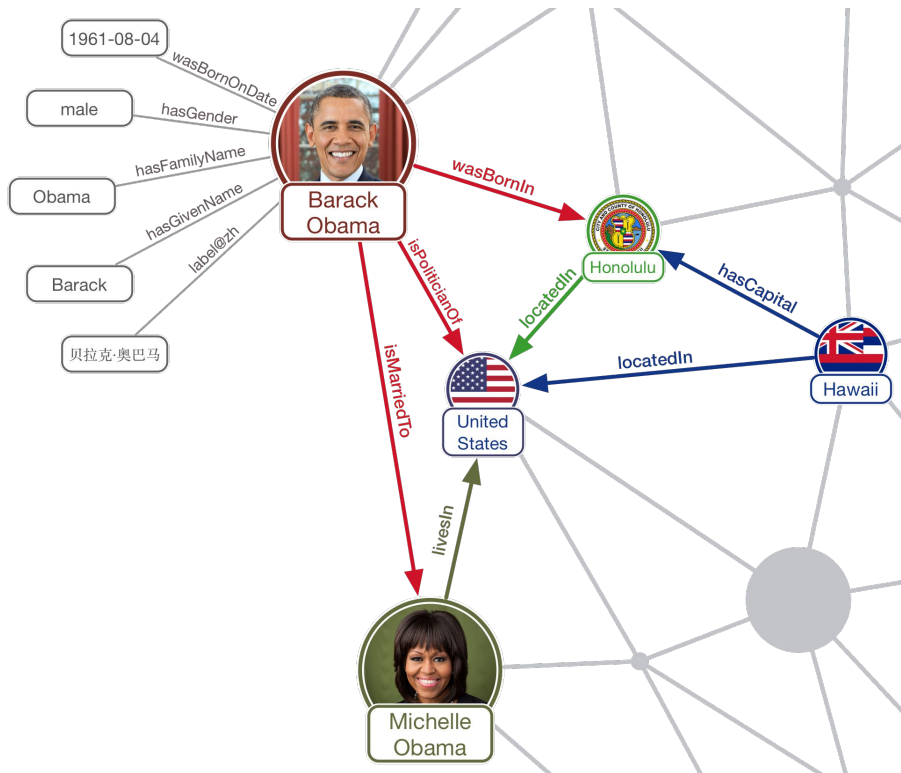# Multi-level Representations *for* Fine-Grained Typing *of* Knowledge Base Entities
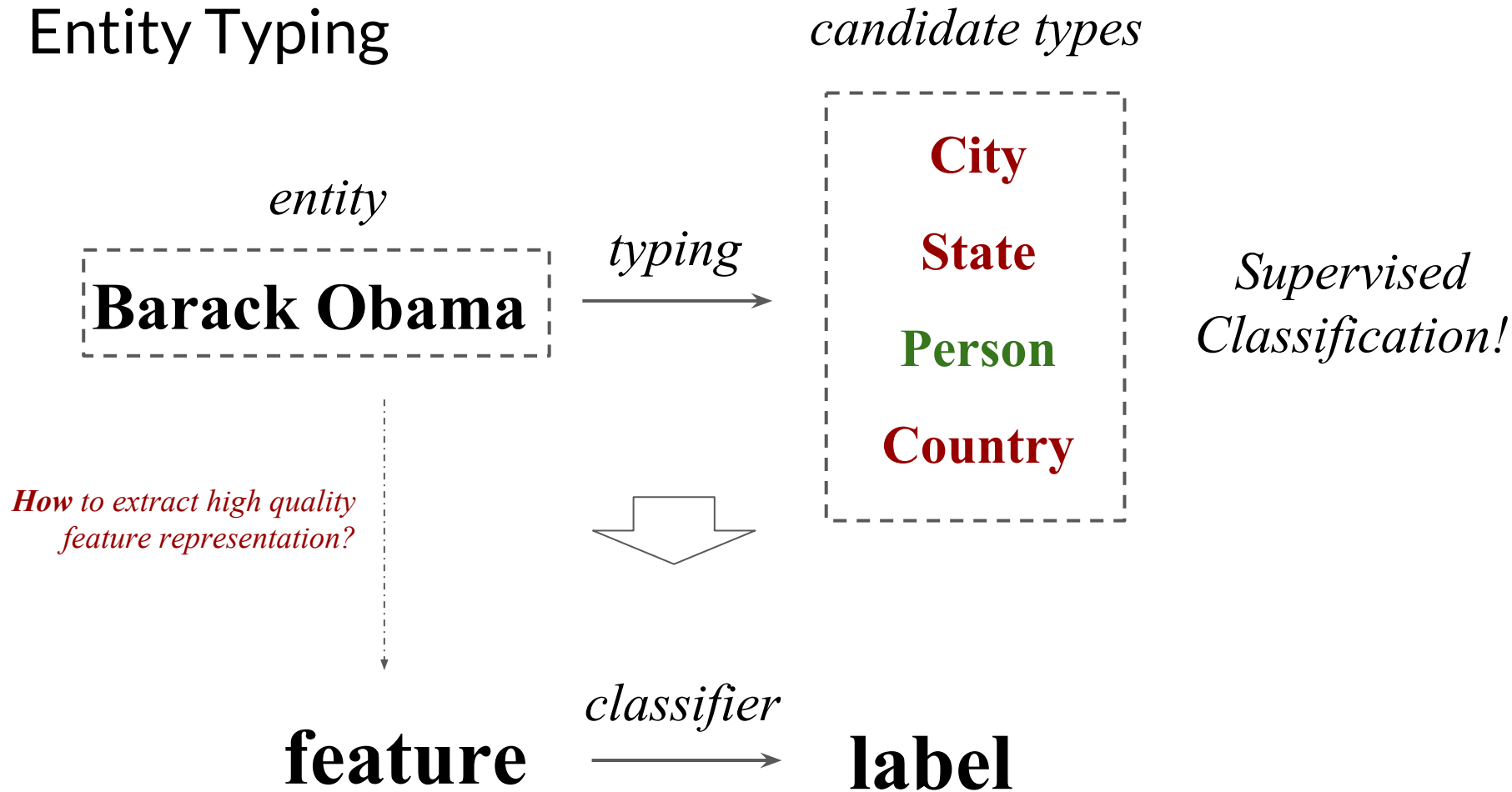
Yadollah Yaghoobzadeh and Hinrich Schütze
LMU Munich, Germany

Presented by: Xiaotao Gu

# Knowledge Graph/Base

# Entity Typing

# Joint Representation

# Character-level Representation



Character-level Representation

Max Pooling

concatenate

Convolution Layer

Filter of size 2
(3 * 6 feature map)

Filter of size 4
(4 * 4 feature map)

Character Embedding of "S"

Character Embedding + Lookup Table

s p a n i s h
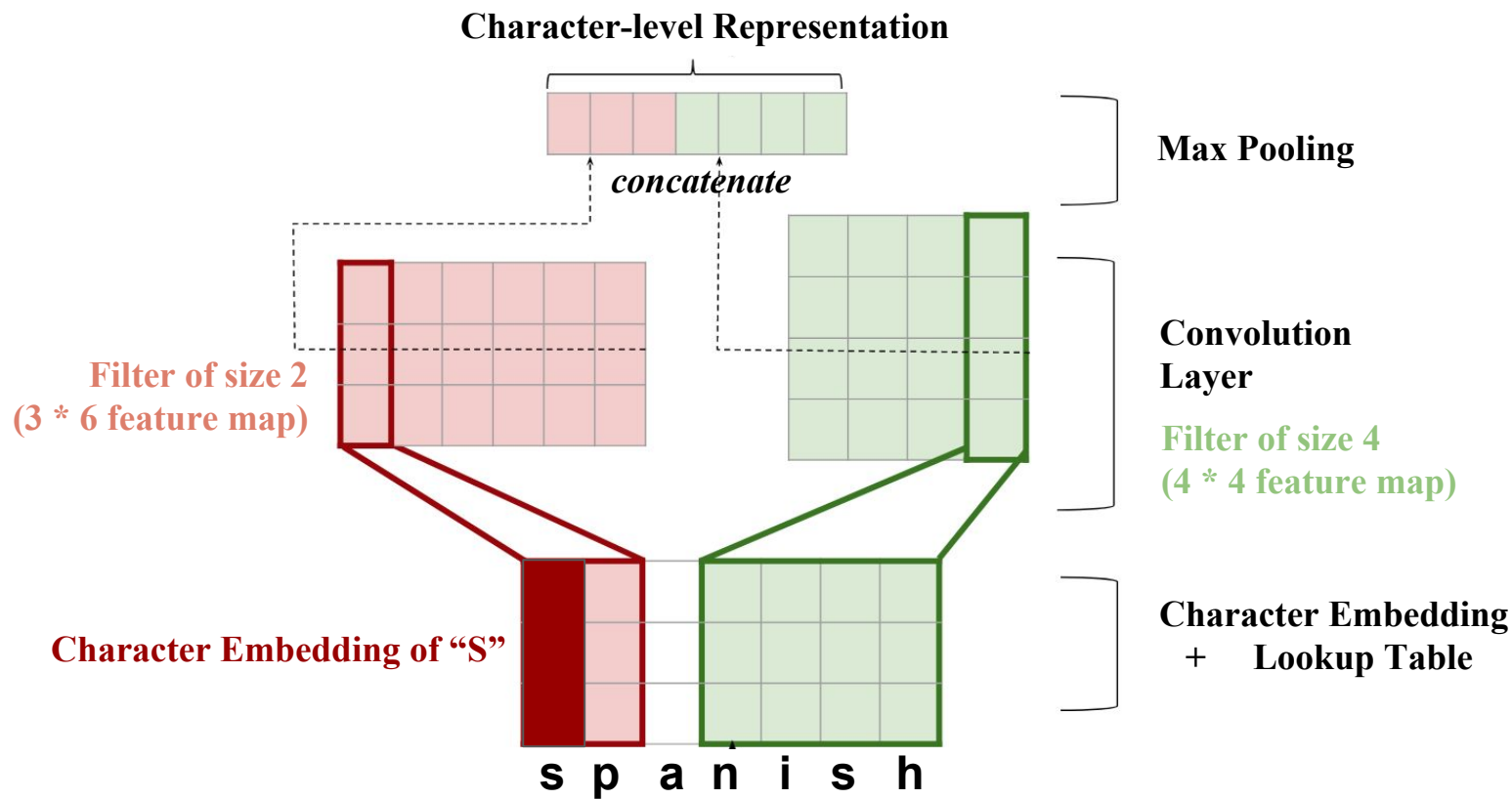
# Joint Representation

# Word-level Representation

Word embedding based on
***Distributional Hypothesis***

Semantic meaning of words are useful for typing!

E.g. "XXX **Lake**" implies **$LOCATION**

Entity embedding = average word embedding

$$E(\textbf{\textit{Lake Michigan}}) = 0.5 * \{ E(\textbf{\textit{Lake}}) + E(\textbf{\textit{Michigan}}) \}$$

Subword/morphology in words are useful for typing!

E.g. "Span**ish**" implies **$LANGUAGE**

***FastText***: n-gram subword embedding!

FastText: https://research.fb.com/fasttext/

## What is *tezgüino* ?

A bottle of tezgüino is on the table.
Everybody likes tezgüino.
Tezgüino makes you drunk.
We make tezgüino out of corn.
(Lin, 1998; Nida, 1975)

## Distributional hypothesis:
You shall know a word by the company it keeps.
(Firth 1957)

The contexts in which a word appears
tells us a lot about what it means.

CS447: Natural Language Processing (J. Hockenmaier)

Image retrieved from:
Professor Julia Hockenmaier's slides for CS447: Natural Language Processing

# Joint Representation

# Entity-level Representation

A bottle of tezgüino is on the table.
Everybody likes tezgüino.
Tezgüino makes you drunk.
We make tezgüino out of corn.

*Unique Identifier*

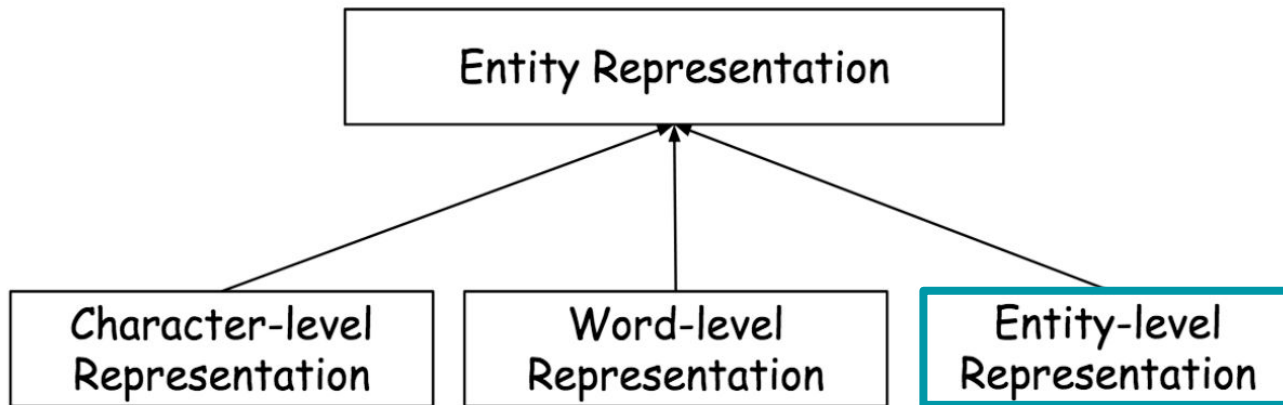*Corpus+SkipGram*

Entity (***Lake Michigan***) ⇒ Identifier (***Lake_Michigan***) ⇒ Entity Embedding

⇓

*Capture **Context** Information*
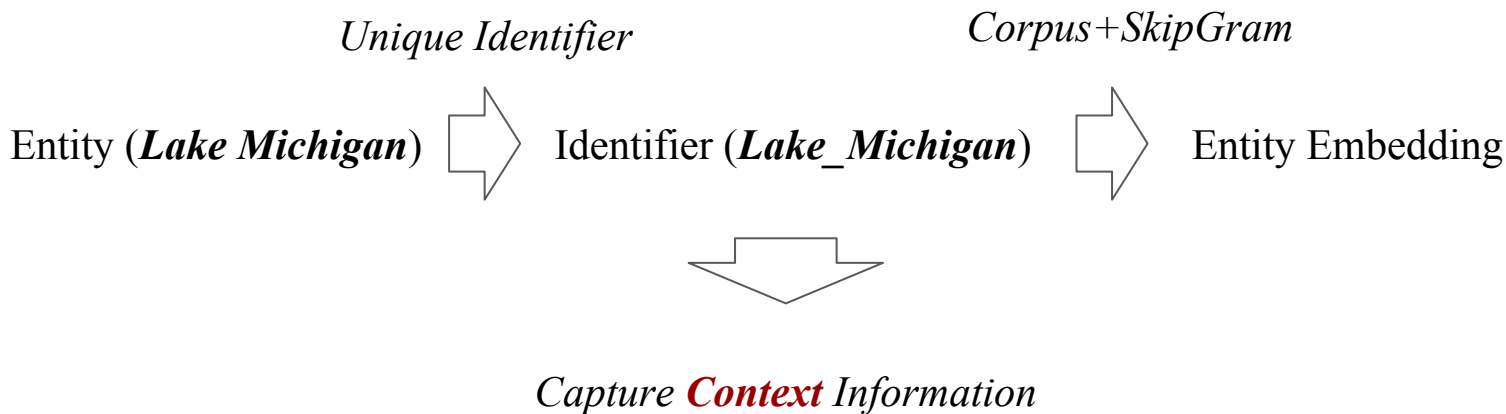
# Entity-level Representation
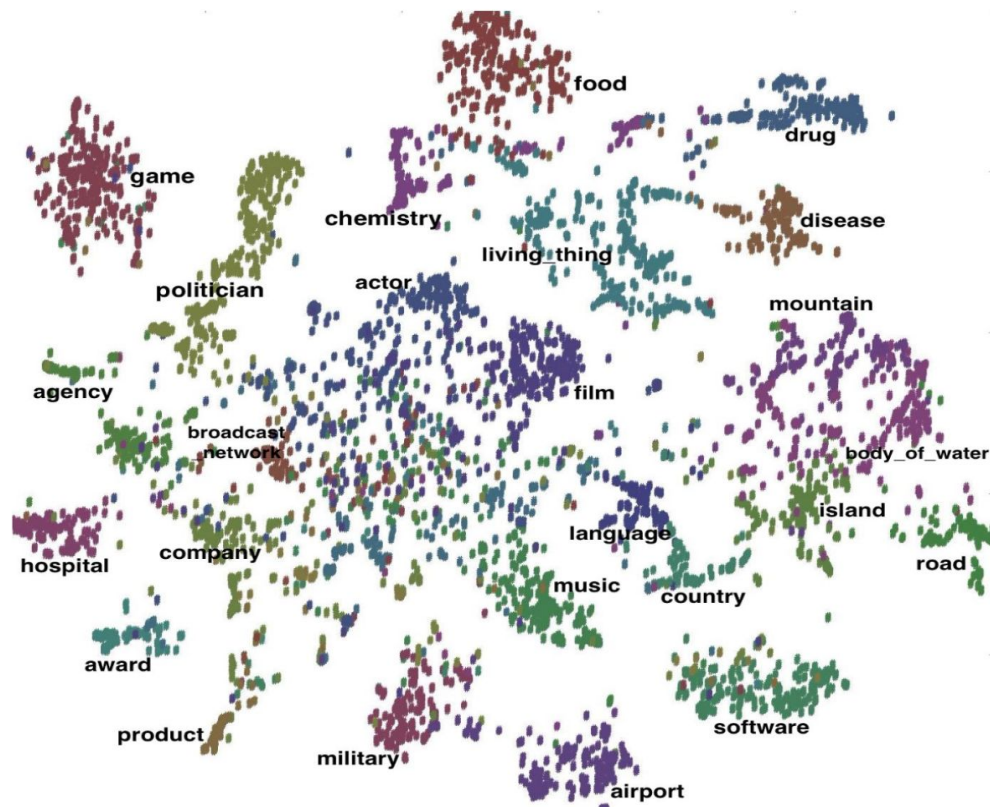
# Joint Representation
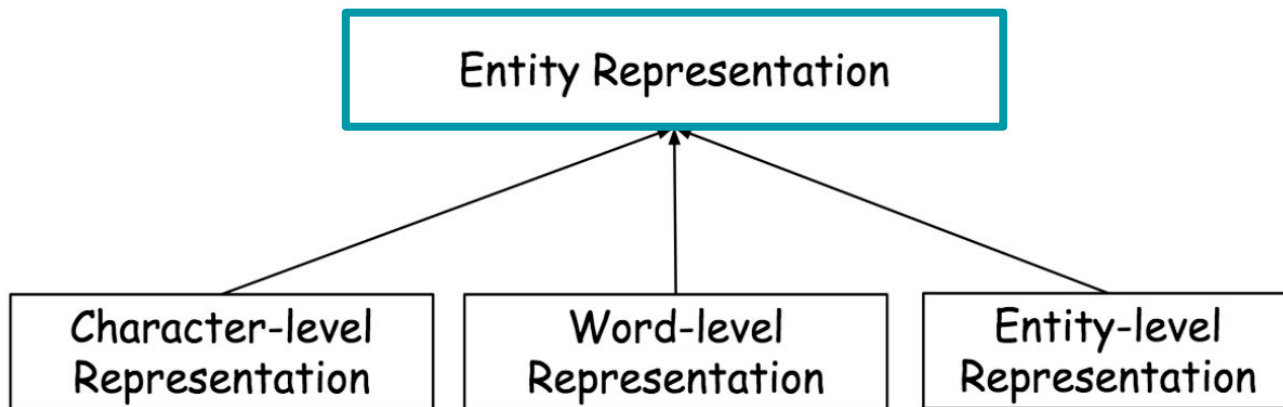
# Experiment

Task: Entity Typing

Dataset: FIGMENT

- 102 types

- 200K Freebase entities, 60K for testing

- 12K head entities (freq > 100), 10K tail entities (freq < 5)

Evaluation Metrics

- *Accuracy*: correct iff all types of an entity are inferred correct and no wrong types are inferred

- *Micro* average F1: F1 of all type-entity assignment decisions

- *Macro* average F1: F1 of types assigned to an entity, averaged over entities

# Experiment

|   |   | all entities | | | head entities | | | tail entities | | |
|---|---|---|---|---|---|---|---|---|---|---|
|   |   | acc | mic | mac | acc | mic | mac | acc | mic | mac |
| 1 | MFT | .000 | .041 | .041 | .000 | .044 | .044 | .000 | .038 | .038 |
| 2 | CLR(FORWARD) | .066 | .379 | .352 | .067 | .342 | .369 | .061 | .374 | .350 |
| 3 | CLR(LSTM) | .121 | .425 | .396 | .122 | .433 | .390 | .116 | .408 | .391 |
| 4 | CLR(BiLSTM) | .133 | .440 | .404 | .129 | .443 | .394 | .135 | .428 | .404 |
| 5 | CLR(NSL) | .164 | .484 | .464 | .157 | .470 | .443 | .173 | .483 | .472 |
| 6 | CLR(CNN) | .177 | .494 | .468 | .171 | .484 | .450 | .187 | .489 | .474 |

**baseline** — row 1

**Character-level** — rows 2–6

# Experiment

|  |  | all entities | | | head entities | | | tail entities | | |
|---|---|---|---|---|---|---|---|---|---|---|
|  |  | acc | mic | mac | acc | mic | mac | acc | mic | mac |
| **baseline** | 1 MFT | .000 | .041 | .041 | .000 | .044 | .044 | .000 | .038 | .038 |
| **Character-level** | 2 CLR(FORWARD) | .066 | .379 | .352 | .067 | .342 | .369 | .061 | .374 | .350 |
| | 3 CLR(LSTM) | .121 | .425 | .396 | .122 | .433 | .390 | .116 | .408 | .391 |
| | 4 CLR(BiLSTM) | .133 | .440 | .404 | .129 | .443 | .394 | .135 | .428 | .404 |
| | 5 CLR(NSL) | .164 | .484 | .464 | .157 | .470 | .443 | .173 | .483 | .472 |
| | 6 CLR(CNN) | .177 | .494 | .468 | .171 | .484 | .450 | .187 | .489 | .474 |
| **Word-level** | 7 BOW | .113 | .346 | .379 | .109 | .323 | .353 | .120 | .356 | .396 |
| | 8 WWLR(SKIP) | .214 | .581 | .531 | .293 | .660 | .634 | .173 | .528 | .478 |
| | 9 WWLR(SSKIP) | .223 | .584 | .543 | .306 | .667 | .642 | .183 | .533 | .494 |
| | 10 SWLR | .236 | .590 | .554 | .301 | .665 | .632 | .209 | .551 | .522 |

# Experiment

| | | all entities | | | head entities | | | tail entities | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | acc | mic | mac | acc | mic | mac | acc | mic | mac |
| **baseline** | 1 MFT | .000 | .041 | .041 | .000 | .044 | .044 | .000 | .038 | .038 |
| | 2 CLR(FORWARD) | .066 | .379 | .352 | .067 | .342 | .369 | .061 | .374 | .350 |
| | 3 CLR(LSTM) | .121 | .425 | .396 | .122 | .433 | .390 | .116 | .408 | .391 |
| **Character-level** | 4 CLR(BiLSTM) | .133 | .440 | .404 | .129 | .443 | .394 | .135 | .428 | .404 |
| | 5 CLR(NSL) | .164 | .484 | .464 | .157 | .470 | .443 | .173 | .483 | .472 |
| | 6 CLR(CNN) | .177 | .494 | .468 | .171 | .484 | .450 | .187 | .489 | .474 |
| | 7 BOW | .113 | .346 | .379 | .109 | .323 | .353 | .120 | .356 | .396 |
| **Word-level** | 8 WWLR(SKIP) | .214 | .581 | .531 | .293 | .660 | .634 | .173 | .528 | .478 |
| | 9 WWLR(SSKIP) | .223 | .584 | .543 | .306 | .667 | .642 | .183 | .533 | .494 |
| | 10 SWLR | .236 | .590 | .554 | .301 | .665 | .632 | .209 | .551 | .522 |
| **Word + Character** | 11 BOW+CLR(NSL) | .156 | .487 | .464 | .157 | .480 | .452 | .159 | .485 | .469 |
| | 12 WWLR+CLR(CNN) | .257 | .603 | .568 | .317 | .668 | .637 | .235 | .567 | .538 |
| | 13 SWLR+CLR(CNN) | .241 | .594 | .561 | .295 | .659 | .628 | .227 | .560 | .536 |

# Experiment

|  |  | all entities | | | head entities | | | tail entities | | |
|---|---|---|---|---|---|---|---|---|---|---|
|  |  | acc | mic | mac | acc | mic | mac | acc | mic | mac |
| **baseline** | 1 MFT | .000 | .041 | .041 | .000 | .044 | .044 | .000 | .038 | .038 |
| | 2 CLR(FORWARD) | .066 | .379 | .352 | .067 | .342 | .369 | .061 | .374 | .350 |
| | 3 CLR(LSTM) | .121 | .425 | .396 | .122 | .433 | .390 | .116 | .408 | .391 |
| **Character-level** | 4 CLR(BiLSTM) | .133 | .440 | .404 | .129 | .443 | .394 | .135 | .428 | .404 |
| | 5 CLR(NSL) | .164 | .484 | .464 | .157 | .470 | .443 | .173 | .483 | .472 |
| | 6 CLR(CNN) | .177 | .494 | .468 | .171 | .484 | .450 | .187 | .489 | .474 |
| | 7 BOW | .113 | .346 | .379 | .109 | .323 | .353 | .120 | .356 | .396 |
| **Word-level** | 8 WWLR(SKIP) | .214 | .581 | .531 | .293 | .660 | .634 | .173 | .528 | .478 |
| | 9 WWLR(SSKIP) | .223 | .584 | .543 | .306 | .667 | .642 | .183 | .533 | .494 |
| | 10 SWLR | .236 | .590 | .554 | .301 | .665 | .632 | .209 | .551 | .522 |
| **Word + Character** | 11 BOW+CLR(NSL) | .156 | .487 | .464 | .157 | .480 | .452 | .159 | .485 | .469 |
| | 12 WWLR+CLR(CNN) | .257 | .603 | .568 | .317 | .668 | .637 | .235 | .567 | .538 |
| | 13 SWLR+CLR(CNN) | .241 | .594 | .561 | .295 | .659 | .628 | .227 | .560 | .536 |
| **Entity-level** | 14 ELR(SKIP) | .488 | .774 | .741 | .551 | .834 | .815 | .337 | .621 | .560 |
| | 15 ELR(SSKIP) | .515 | .796 | .763 | .560 | .839 | .819 | .394 | .677 | .619 |

# Experiment

|  |  | all entities | | | head entities | | | tail entities | | |
|---|---|---|---|---|---|---|---|---|---|---|
|  |  | acc | mic | mac | acc | mic | mac | acc | mic | mac |
| **baseline** | 1 MFT | .000 | .041 | .041 | .000 | .044 | .044 | .000 | .038 | .038 |
| | 2 CLR(FORWARD) | .066 | .379 | .352 | .067 | .342 | .369 | .061 | .374 | .350 |
| **Character-level** | 3 CLR(LSTM) | .121 | .425 | .396 | .122 | .433 | .390 | .116 | .408 | .391 |
| | 4 CLR(BiLSTM) | .133 | .440 | .404 | .129 | .443 | .394 | .135 | .428 | .404 |
| | 5 CLR(NSL) | .164 | .484 | .464 | .157 | .470 | .443 | .173 | .483 | .472 |
| | 6 CLR(CNN) | .177 | .494 | .468 | .171 | .484 | .450 | .187 | .489 | .474 |
| | 7 BOW | .113 | .346 | .379 | .109 | .323 | .353 | .120 | .356 | .396 |
| **Word-level** | 8 WWLR(SKIP) | .214 | .581 | .531 | .293 | .660 | .634 | .173 | .528 | .478 |
| | 9 WWLR(SSKIP) | .223 | .584 | .543 | .306 | .667 | .642 | .183 | .533 | .494 |
| | 10 SWLR | .236 | .590 | .554 | .301 | .665 | .632 | .209 | .551 | .522 |
| **Word + Character** | 11 BOW+CLR(NSL) | .156 | .487 | .464 | .157 | .480 | .452 | .159 | .485 | .469 |
| | 12 WWLR+CLR(CNN) | .257 | .603 | .568 | .317 | .668 | .637 | .235 | .567 | .538 |
| | 13 SWLR+CLR(CNN) | .241 | .594 | .561 | .295 | .659 | .628 | .227 | .560 | .536 |
| **Entity-level** | 14 ELR(SKIP) | .488 | .774 | .741 | .551 | .834 | .815 | .337 | .621 | .560 |
| | 15 ELR(SSKIP) | .515 | .796 | .763 | .560 | .839 | .819 | .394 | .677 | .619 |
| | 16 AGG-FIGER | .320 | .694 | .660 | .396 | .762 | .724 | .220 | .593 | .568 |
| **Entity + Word + Character** | 17 ELR+CLR | .554 | .816 | .788 | .580 | .844 | .825 | .467 | .733 | .690 |
| | 18 ELR+WWLR | .557 | .819 | .793 | .582 | .846 | .827 | .480 | .749 | .708 |
| | 19 ELR+SWLR | .558 | .820 | .796 | .584 | .846 | .829 | .480 | .751 | .714 |
| | 20 ELR+WWLR+CLR | .568 | .823 | .798 | .590 | .847 | .829 | .491 | .755 | .716 |
| | 21 ELR+SWLR+CLR | .569 | .824 | .801 | .590 | .849 | **.831** | .497 | .760 | .724 |
| | 22 ELR+WWLR+CLR+TC | .572 | .824 | .801 | .594 | .849 | **.831** | .499 | .759 | .722 |
| | 23 ELR+SWLR+CLR+TC | **.575** | **.826** | **.802** | **.597** | **.851** | **.831** | **.508** | **.762** | **.727** |

# Experiment

Character-Level: **CNN** performs best for capturing local features

Word-Level: **Subword** improves the performance, especially for rare entities

**Joint** entity-word-character information: achieves the best performance

External Information: adding **description** text helps tail entities!

| entities: | all | head | tail |
|---|---|---|---|
| AVG-DES | .773 | .791 | .745 |
| MuLR | .825 | .846 | .757 |
| MuLR+AVG-DES | **.873** | **.877** | **.852** |

Add description

# Summary

High-quality Entity Representation from:

- Context information: from large corpus

- Surface name information: word, subword, character-level information

- External knowledge: description, relational links

- The joint representation is the most informative for entity typing.

***Thank you!***